

# A small delay and correlation time limit of stochastic differential delay equations with state-dependent colored noise

Scott Hottoy<sup>1</sup>, Austin McDaniel<sup>2,3</sup>, and Jan Wehr<sup>2</sup>

<sup>1</sup>Department of Mathematics, University of Wisconsin-Madison,  
Madison, Wisconsin 53706 USA

<sup>2</sup>Department of Mathematics, University of Arizona, Tucson,  
Arizona 85721 USA

<sup>3</sup>School of Mathematical and Statistical Sciences, Arizona State  
University, Tempe, Arizona 85287 USA

## Abstract

We consider a general stochastic differential delay equation (SDDE) with state-dependent colored noises and derive its limit as the time delays and the correlation times of the noises go to zero. The work is motivated by an experiment involving an electrical circuit with noisy, delayed feedback. In previous work, the delay equation is approximated by using a Taylor expansion. Here we derive the limit without performing such an approximation and thereby get a different and more accurate result. An Ornstein-Uhlenbeck process is used to model the colored noise. The main methods used in the proof are a theorem about convergence of stochastic integrals by Kurtz and Protter and a maximal inequality for sums of a stationary sequence of random variables by Peligrad and Utev. The limiting equation that we obtain may be used as a working SDE approximation of an SDDE modeling a real system where noises are correlated in time and whose response to noise sources depends on the system's state at a previous time.

## 1 Introduction

Stochastic differential equations (SDEs) are frequently used to describe the dynamics of physical and biological systems [12]. However, in situations where a system's response to stimuli is delayed, stochastic differential delay equations (SDDEs) provide more accurate models. This gain in accuracy comes at a price

of greater mathematical difficulty because the theory of SDDEs is much less developed than the theory of SDEs. Thus, it is useful to develop approximations of SDDEs that are easier to work with than the original equations but still account for the effects of the delay(s).

In this article we consider a general SDDE system, driven by colored (i.e. temporally correlated) noise, which was motivated by an experiment involving an electrical circuit with a delayed feedback mechanism [16]. In the experiment, the voltage is driven by a colored noise process with a rapidly decaying correlation function. An SDE approximation of an SDDE that generalizes the circuit was derived by first performing a Taylor expansion to first order in the time delays and then taking the limit in which the time delays and the correlation times of the noises go to zero at the same rate [16]. This limiting SDE contains *noise-induced drift* terms which depend on the ratios of the time delays to the noise correlation times. Convergence of the solution of the equation obtained by Taylor expansion to the solution of the limiting equation was proven later in [10].

The present article improves upon the approximation contained in [10, 16]. We study the same limit of the SDDE system, but without first performing a Taylor approximation, and thereby get a more accurate result. Because we do not perform a Taylor expansion, we are able to use a simpler (less smooth) model of the colored noise process than the one used in [10, 16]; here, we model the colored noise as a stationary solution to an Ornstein-Uhlenbeck SDE [4]. Our refinement of the results of [10, 16] can be used in practice as an SDE approximation of a general class of systems with delay.

We consider the multidimensional SDDE system

$$d\mathbf{x}_t = \mathbf{f}(\mathbf{x}_t)dt + \boldsymbol{\sigma}(\mathbf{x}_{t-\boldsymbol{\delta}})\boldsymbol{\xi}_t dt \quad (1)$$

where  $\mathbf{x}_t \in \mathbb{R}^m$  is the state vector,  $\mathbf{f}(\mathbf{x}) \in \mathbb{R}^m$  is a vector-valued function describing the deterministic part of the dynamical system,  $\boldsymbol{\xi}_t \in \mathbb{R}^n$  is a vector of zero-mean independent noises,  $\boldsymbol{\sigma}(\mathbf{x}) \in \mathbb{R}^{m \times n}$  is a matrix-valued function, and  $\mathbf{x}_{t-\boldsymbol{\delta}} = ((x_{t-\delta_1})_1, \dots, (x_{t-\delta_i})_i, \dots, (x_{t-\delta_m})_m)^T$  (where T denotes transpose) is the delayed state vector. Note that each component is delayed by a possibly different amount  $\delta_i > 0$ . Each of the  $n$  independent noises  $\xi_j$  is colored and therefore characterized by a correlation time  $\tau_j > 0$ . That is, assuming stationarity,  $E[(\xi_t)_j(\xi_s)_j] = g(|t-s|\tau_j^{-1})$  where  $g$  is a function that decays quickly as its argument increases (for  $i \neq j$ ,  $E[(\xi_t)_i(\xi_s)_j] = 0$  by independence).

In the main theorem of the article (Theorem 1), we consider the case where the components of  $\boldsymbol{\xi}$  are independent Ornstein-Uhlenbeck colored noises with correlation times  $\tau_j$ . That is, we define  $\xi_j = y_j$  where  $y_j$  is the solution of

$$d(y_t)_j = -\frac{1}{\tau_j}(y_t)_j dt + \frac{1}{\tau_j}d(W_t)_j \quad (2)$$

where  $\tau_j > 0$  and  $\mathbf{W}_t \in \mathbb{R}^n$  is an  $n$ -dimensional Wiener process. Equation (2) has a unique stationary measure and the distribution of  $\mathbf{y}_t$  with an arbitrary initial condition converges to this stationary measure as  $t \rightarrow \infty$ . The solution of

(2), with the initial condition distributed according to the stationary measure, defines a stationary process whose realizations will play the role of colored noise in the SDDE system (1). Note that as  $\tau_j \rightarrow 0$  for all  $j = 1, \dots, n$ ,  $\mathbf{y}$  converges to an  $n$ -dimensional white noise (i.e., its correlation function converges to a delta function).

We study the limit of the system consisting of equations (1) and (2), with  $\xi_t = \mathbf{y}_t$ , assuming that all  $\delta_i$  and  $\tau_j$  stay proportional to a single characteristic time  $\epsilon > 0$  which goes to 0. Thus, we let  $\delta_i = c_i \epsilon \rightarrow 0$  and  $\tau_j = k_j \epsilon \rightarrow 0$  where  $c_i, k_j > 0$  remain constant for all  $i, j$  and  $\epsilon \rightarrow 0$ .

We consider the solution to equations (1) and (2) on a bounded time interval  $0 \leq t \leq T$ . We let  $(\Omega, \mathcal{F}, P)$  denote the underlying probability space. In order to formulate a well-posed problem, because of the delays in (1), one needs to specify not only an initial condition but also the values of the process  $\mathbf{x}$  at all past times  $t \in [-(\max_i \delta_i), 0]$ . We will require that values of  $\mathbf{x}$  are initially specified on the larger interval  $t \in [-2(\max_i \delta_i), 0]$  and also that values of  $\mathbf{y}$  are specified on the interval  $t \in [-(\max_i \delta_i), 0]$ . In other words, we assume that there is a  $t_- < 0$  such that the values of  $\mathbf{x}$  and  $\mathbf{y}$  are initially specified for  $t \in [t_-, 0]$  and we only consider delays  $\delta_i$  such that  $\delta_i < |t_-|/2$  for all  $i$ . Let  $\mathbf{x}^- : \Omega \times [t_-, 0] \rightarrow \mathbb{R}^m$  denote the *past condition* associated with (1). We assume that  $\mathbf{x}^-$  is independent of  $\mathbf{W}_t$  for all  $t \geq 0$ . We further assume that  $\mathbf{x}^-$  is defined so that there exists a unique solution to (1) with the past condition  $\mathbf{x}^-$ . Similarly, let  $\mathbf{y}^- : \Omega \times [t_-, 0] \rightarrow \mathbb{R}^n$  specify the values of  $\mathbf{y}$  on the interval  $[t_-, 0]$ .

**Theorem 1.** *Suppose that  $\mathbf{f}$  is continuous and bounded and that  $\boldsymbol{\sigma}$  is bounded with bounded, Lipschitz continuous first derivatives. Let  $(\mathbf{x}^\epsilon, \mathbf{y}^\epsilon)$  solve equations (1) and (2) (which depend on  $\epsilon$  through  $\delta_i, \tau_j$ ) on  $0 \leq t \leq T$  with the past condition  $\mathbf{x}^-$  the same for every  $\epsilon$  and the initial condition  $\mathbf{y}_0^\epsilon = (\mathbf{y}^-)_0^\epsilon$  where  $(\mathbf{y}^-)^\epsilon$  is defined so that  $\mathbf{y}^\epsilon$  considered on the interval  $t > t_-$  is a stationary process. Let  $\mathbf{x}$  solve*

$$d\mathbf{x}_t = \mathbf{f}(\mathbf{x}_t)dt + \mathbf{S}(\mathbf{x}_t) dt + \boldsymbol{\sigma}(\mathbf{x}_t)d\mathbf{W}_t \quad (3)$$

on  $0 \leq t \leq T$  with the same initial condition  $\mathbf{x}_0 = \mathbf{x}_0^-$ , where  $\mathbf{S}$  is defined componentwise as

$$S_i(\mathbf{x}) = \sum_{p,j} \frac{1}{2} e^{-\frac{\delta_p}{\tau_j}} \sigma_{pj}(\mathbf{x}) \frac{\partial \sigma_{ij}(\mathbf{x})}{\partial x_p} \quad (4)$$

and suppose strong uniqueness holds on  $0 \leq t \leq T$  for (3) with the initial condition  $\mathbf{x}_0$ . Then

$$\lim_{\epsilon \rightarrow 0} P \left[ \sup_{0 \leq t \leq T} \|\mathbf{x}_t^\epsilon - \mathbf{x}_t\| > a \right] = 0 \quad (5)$$

for every  $a > 0$ .

An outline of the paper is as follows. In Section 2 we discuss the important modeling aspects of a general colored noise process  $\xi$ . In Section 3 we prove the main theorem of the paper. In Section 4, we discuss how the theorem relates to the previous works [10, 16]. We give conclusions in Section 5.

## 2 Colored noise process

In this section we discuss the specific model of colored noise that we use in this paper, i.e., the Ornstein-Uhlenbeck (OU) process. The noise process  $\xi$  driving the system (1) is colored, not white. The terms “colored” and “white” come from the Wiener-Khinchin theorem (see [4, Section 1.5.2]). This theorem states that the expected value of the modulus of the Fourier transform of the time series of the noise is equal to the Fourier transform of its time correlation function. Thus, a white noise process has a constant frequency correlation function in the Fourier domain (much like white light contains all colors of the light spectrum in equal proportions). A colored noise is any (usually mean zero) process  $\xi$  with a nonconstant frequency correlation function. In this paper, we are interested in colored noise processes  $\xi$  which have a time correlation function of the form

$$E[\xi_t \xi_s] = \frac{1}{\epsilon} g\left(\frac{|t-s|}{\epsilon}\right)$$

where  $g$  is a function that decays rapidly as its argument increases. For small  $\epsilon > 0$ , the correlation function can be approximated by the correlation function of white noise, i.e.,

$$E[\xi_t \xi_s] \rightarrow \delta(t-s),$$

as  $\epsilon \rightarrow 0$ . Furthermore, it is typical to use as a model for colored noise a process  $\xi$  such that

$$\int_0^t \xi_s ds \rightarrow W_t,$$

as  $\epsilon \rightarrow 0$ , in some probabilistic sense.

Once the correlation function  $g(r)$ ,  $r \geq 0$ , (or color) of the noise process is chosen, the smoothness of the noise should be determined. Different processes with varying degrees of smoothness have been used for modeling colored noise in the literature. An infinitely differentiable process was used in [3] by convoluting a smooth function  $h \in C^\infty([0, \infty))$  with a Wiener process. A piecewise differentiable approximation of a Wiener process was constructed in [9] which was then differentiated to obtain a colored noise. In [10, 16], a differentiable harmonic noise process was used in order to make the solution twice differentiable and thus allow to use its Taylor expansion.

In this paper, we use an Ornstein-Uhlenbeck (OU) process to model the colored noise. This is an often used simple choice (see, e.g., [6, 8, 13, 14]) because it is a continuous process that can be written as the solution to a linear SDE. An OU process  $y$  with paths in  $C([0, T], \mathbb{R})$  is defined as the solution to the SDE

$$dy_t = -\alpha y_t dt + \sigma dW_t, \quad y_0 = w, \quad (6)$$

where  $\alpha, \sigma > 0$ . This process  $y$  is a Gaussian process with mean  $E[y_t] = e^{-\alpha t} E[w]$  and covariance [4, Section 4.5.4]

$$E[(y_t - E[y_t])(y_s - E[y_s])] = \left[ \text{Var}(w) - \frac{\sigma^2}{2\alpha} \right] e^{-\alpha(t+s)} + \frac{\sigma^2}{2\alpha} e^{-\alpha|t-s|}. \quad (7)$$

Furthermore, the OU process is ergodic, that is, it has a unique stationary measure, such that starting from any initial distribution, the system's distribution converges to this measure. In Section 1, we defined the colored noises to be  $n$  independent, stationary OU processes with correlation times  $\tau_j = k_j \epsilon$ . In other words,  $(\xi_t)_j = (y_t^\epsilon)_j$  is the solution to (6) with  $\alpha = \sigma = \tau_j^{-1}$  and with  $(y_0^\epsilon)_j = w_j$  where  $w_j$  is a random variable, independent of the process  $W$ , having this stationary distribution, i.e.,  $w_j$  is Gaussian with mean zero and variance  $1/(2k_j\epsilon)$ .

The choice of this particular colored noise is advantageous because equation (6) is exactly solvable and every moment of  $\mathbf{y}_t^\epsilon$  can be calculated. Furthermore, the covariance moments  $E[(y_t^\epsilon)_j (y_t^\epsilon)_\ell]^n$  can be calculated for all  $n \in \mathbb{N}$  by using Wick's theorem [7, Theorem 1.3.8].

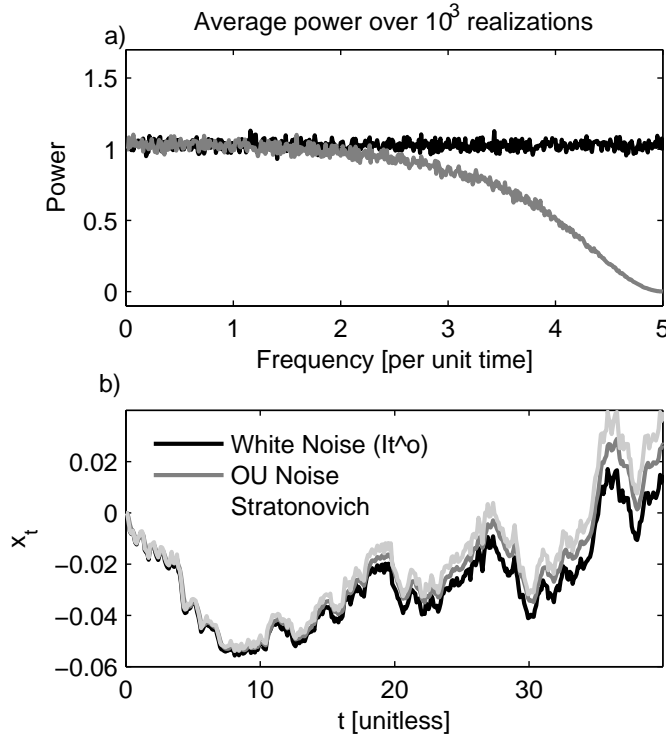


Figure 1: Comparison of white and Ornstein-Uhlenbeck noises. Plot a) is the average power, or modulus of the Fourier time series, of each noise. Plot b) is a realization of the solution to system (1) with  $\delta = 0$ ,  $f(x) = ax$ ,  $\sigma(x) = (bx + c)$ , and the initial condition  $x_0 = 0$ , where  $\xi_t$  is each type of noise.

In Figure 1, white noise is compared to OU noise. White noise (black) is generated by taking the differential of a standard Wiener process. Ornstein-

Uhlenbeck noise (dark gray) is generated by using the SDE (2) with  $\tau = 5$  (arbitrary time units). In panel a) the average power as a function of frequency is plotted for each process. This is the Fourier transform pair of the correlation function. For the white noise process the average power is constant as a function of frequency, while the average power of the OU process decays rapidly after frequency 2. In panel b), realizations are shown of a one-dimensional example of system (1), with no delay, driven by the noises. Notice that the processes driven by the OU noise and the white noise diverge from each other after time  $t = 20$ . This is a result of the Stratonovich correction that appears in (3) when one sets the delay equal to zero.

Theorem 1, as stated, only applies to the case where the noise is modeled by a stationary OU process. The theorem can be modified to handle any noise process that is stationary and that solves a linear SDE with additive (white) noise (e.g. the harmonic noise process in [10, 16]). For different noises, the coefficients of the additional drift in the analogous limiting equation depend on the covariance function of the noise. We expect that, with extra work, stationary noises defined by other SDEs may also be treated. For a general noise  $\xi$ , defined by its covariance function, different methods must be employed to prove an analogous theorem.

### 3 Proof of Theorem 1

In this section we prove the main theorem of the paper, Theorem 1. The main tool that we use is a theory of convergence of stochastic integrals by Kurtz and Protter [9]. A similar technique is used in [5] and in [10]. The structure of the section is as follows. In Section 3.1 we introduce the theory of convergence of stochastic integrals that we will use to prove the theorem. In Section 3.2 we use integration by parts and substitution to write system (1) in the form necessary for applying the Kurtz-Protter theorem. In Section 3.3 we complete the proof of Theorem 1 by verifying that the conditions of the Kurtz-Protter theorem are satisfied.

We begin with the theory of convergence of stochastic integrals, where we state (in a less general but sufficient form) a theorem of Kurtz and Protter [9].

#### 3.1 Convergence of stochastic integrals

We fix a probability space  $(\Omega, \mathcal{F}, P)$  and an  $n$ -dimensional Wiener process  $W$  on it. Let  $\mathcal{F}_t$  will be (the usual augmentation of)  $\sigma(\{W_s : s \leq t\})$ , the filtration generated by  $W$  up to time  $t$  [18].

Suppose  $H$  is an  $\{\mathcal{F}_t\}$ -adapted semimartingale with paths in  $C([0, T], \mathbb{R}^k)$ , whose Doob-Meyer decomposition is  $H_t = M_t + A_t$  so that  $M$  is an  $\mathcal{F}_t$ -local martingale and  $A$  is a process of locally bounded variation, such that  $A_0 = 0$  [18]. For a continuous  $\{\mathcal{F}_t\}$ -adapted process  $Y$  with paths in  $C([0, T], \mathbb{R}^{d \times k})$

and for  $t \leq T$  consider the Itô integral

$$\int_0^t \mathbf{Y}_s d\mathbf{H}_s = \lim \sum_i \mathbf{Y}_{t_i} (\mathbf{H}_{t_{i+1}} - \mathbf{H}_{t_i}) , \quad (8)$$

where  $\{t_i\}$  is a partition of  $[0, t]$  and the limit is taken as the maximum of  $t_{i+1} - t_i$  goes to zero. For a continuous processes  $\mathbf{Y}$  such that

$$P \left( \int_0^T \|\mathbf{Y}_s\|^2 d\langle \mathbf{M} \rangle_s + \int_0^T \|\mathbf{Y}_s\| dV_s(\mathbf{A}) < \infty \right) = 1 ,$$

where  $\langle \mathbf{M} \rangle_s$  is the quadratic variation of  $\mathbf{M}$  and  $V_s(\mathbf{A})$  is the total variation of  $\mathbf{A}$ , the limit in equation (8) exists in the sense that

$$\sup_{0 \leq t \leq T} \left( \left\| \int_0^t \mathbf{Y}_s d\mathbf{H}_s - \sum_i \mathbf{Y}_{t_i} (\mathbf{H}_{t_{i+1}} - \mathbf{H}_{t_i}) \right\| \right) \rightarrow 0$$

in probability. This (and related) convergence modes will be used throughout the paper [17]

Consider  $(\mathbf{U}^\epsilon, \mathbf{H}^\epsilon)$  with paths in  $C([0, T], \mathbb{R}^m \times \mathbb{R}^k)$  adapted to  $\{\mathcal{F}_t\}$  where  $\mathbf{H}^\epsilon$  is a semimartingale with respect to  $\mathcal{F}_t$ . Let  $\mathbf{H}_t^\epsilon = \mathbf{M}_t^\epsilon + \mathbf{A}_t^\epsilon$  be its Doob-Meyer decomposition. Let  $\mathbf{h} : \mathbb{R}^k \rightarrow \mathbb{R}^{k \times n}$  be a continuous matrix-valued function and let  $\mathbf{X}^\epsilon$ , with paths in  $C([0, T], \mathbb{R}^m)$ , satisfy the SDE

$$\mathbf{X}_t^\epsilon = \mathbf{X}_0 + \mathbf{U}_t^\epsilon + \int_0^t \mathbf{h}(\mathbf{X}_s^\epsilon) d\mathbf{H}_s^\epsilon , \quad (9)$$

where  $\mathbf{X}_0^\epsilon = \mathbf{X}_0 \in \mathbb{R}^m$  is the same initial condition for all  $\epsilon$ . Define  $\mathbf{X}$ , with paths in  $C([0, T], \mathbb{R}^m)$ , to be the solution of

$$\mathbf{X}_t = \mathbf{X}_0 + \int_0^t \mathbf{h}(\mathbf{X}_s) d\mathbf{H}_s . \quad (10)$$

Note that (9) implies  $\mathbf{U}_0^\epsilon = \mathbf{0}$  for all  $\epsilon$ .

**Lemma 1.** [9, Theorem 5.10] Suppose  $(\mathbf{U}^\epsilon, \mathbf{H}^\epsilon) \rightarrow (\mathbf{0}, \mathbf{H})$  in probability with respect to  $C([0, T], \mathbb{R}^m \times \mathbb{R}^k)$ , i.e., for all  $a > 0$ ,

$$P \left[ \sup_{0 \leq s \leq T} (\|\mathbf{U}_s^\epsilon\| + \|\mathbf{H}_s^\epsilon - \mathbf{H}_s\|) > a \right] \rightarrow 0 \quad (11)$$

as  $\epsilon \rightarrow 0$ , and the following condition is satisfied:

**Condition 1.** [Tightness condition] For every  $t \in [0, T]$ , the family of total variations evaluated at  $t$ ,  $\{V_t(\mathbf{A}^\epsilon)\}$ , is stochastically bounded, i.e.  $P[V_t(\mathbf{A}^\epsilon) > L] \rightarrow 0$  as  $L \rightarrow \infty$ , uniformly in  $\epsilon$ .

Suppose that there exists a unique global solution to equation (10). Then, as  $\epsilon \rightarrow 0$ ,  $\mathbf{X}^\epsilon$  converges to  $\mathbf{X}$ , the solution of equation (10), in probability with respect to  $C([0, T], \mathbb{R}^m)$ .

### 3.2 Derivation of the limiting equation

*Proof of Theorem 1.* To write system (1) (with  $\boldsymbol{\xi}_t = \mathbf{y}_t^\epsilon$ ) in the form of (9), we will use integration by parts and substitution. To this end, we write equation (2) in matrix form. Recalling that  $\tau_j = k_j \epsilon$  and defining

$$\mathbf{D} = \begin{bmatrix} \frac{1}{k_1} & 0 & \dots & 0 \\ 0 & \frac{1}{k_2} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \frac{1}{k_n} \end{bmatrix},$$

equation (2) becomes

$$d\mathbf{y}_t^\epsilon = \frac{\mathbf{D}}{\epsilon} (-\mathbf{y}_t^\epsilon dt + d\mathbf{W}_t). \quad (12)$$

We solve equation (12) for  $\mathbf{y}_t^\epsilon dt$  and substitute it into equation (1) to obtain

$$d\mathbf{x}_t^\epsilon = \mathbf{f}(\mathbf{x}_t^\epsilon)dt + \boldsymbol{\sigma}(\mathbf{x}_{t-c\epsilon}^\epsilon)d\mathbf{W}_t - \boldsymbol{\sigma}(\mathbf{x}_{t-c\epsilon}^\epsilon)\epsilon\mathbf{D}^{-1}d\mathbf{y}_t^\epsilon$$

where  $\mathbf{x}_{t-c\epsilon}^\epsilon = ((x_{t-c_1\epsilon}^\epsilon)_1, \dots, (x_{t-c_i\epsilon}^\epsilon)_i, \dots, (x_{t-c_m\epsilon}^\epsilon)_m)^T$ . In integral form, this equation becomes

$$\mathbf{x}_t^\epsilon = \mathbf{x}_0 + \int_0^t \mathbf{f}(\mathbf{x}_s^\epsilon)ds + \int_0^t \boldsymbol{\sigma}(\mathbf{x}_{s-c\epsilon}^\epsilon)d\mathbf{W}_s - \int_0^t \boldsymbol{\sigma}(\mathbf{x}_{s-c\epsilon}^\epsilon)\epsilon\mathbf{D}^{-1}d\mathbf{y}_s^\epsilon \quad (13)$$

In the limit as  $\epsilon \rightarrow 0$ , we expect the second and the third terms of the right hand side of equation (13) to converge to the analogous terms of the limiting equation (3). In addition,  $\epsilon\mathbf{y}^\epsilon$  goes to zero as  $\epsilon \rightarrow 0$ , as we will show later (see Lemma 2). Because of this, one might expect the last term of the right hand side to converge to zero as well. This is the case when  $\boldsymbol{\sigma}(\mathbf{x}) = \boldsymbol{\sigma}$  is a constant function. For non-constant  $\boldsymbol{\sigma}(\mathbf{x})$ , in order to be able to apply Lemma 1 directly the process  $\epsilon\mathbf{y}^\epsilon$  would have to satisfy Condition 1. This is not the case, nor is it true for any family of colored noise processes which converge to white noise (see [9]).

To resolve this issue, we integrate by parts the  $i^{\text{th}}$  component of the last integral to obtain

$$\begin{aligned} (x_t^\epsilon)_i &= (x_0)_i + \int_0^t f_i(\mathbf{x}_s^\epsilon)ds + \left( \int_0^t \boldsymbol{\sigma}(\mathbf{x}_{s-c\epsilon}^\epsilon)d\mathbf{W}_s \right)_i - \sum_j \sigma_{ij}(\mathbf{x}_{t-c\epsilon}^\epsilon)k_j\epsilon(y_t^\epsilon)_j \\ &\quad + \sum_j \sigma_{ij}(\mathbf{x}_{t-c\epsilon}^\epsilon)k_j\epsilon(y_0^\epsilon)_j + \int_0^t \sum_{p,j} \frac{\partial}{\partial x_p} [\sigma_{ij}(\mathbf{x}_{s-c\epsilon}^\epsilon)] k_j\epsilon(y_s^\epsilon)_j d(x_{s-c_p\epsilon}^\epsilon)_p \end{aligned}$$



We substitute equation (1) into the term containing  $d(x_{s-c_p\epsilon}^\epsilon)_p$  to obtain

$$\begin{aligned}
(x_t^\epsilon)_i &= (x_0)_i + \int_0^t f_i(\mathbf{x}_s^\epsilon) ds + \left( \int_0^t \boldsymbol{\sigma}(\mathbf{x}_{s-c\epsilon}^\epsilon) d\mathbf{W}_s \right)_i \\
&\quad + \int_0^t \sum_{p,j} \frac{\partial}{\partial x_p} [\sigma_{ij}(\mathbf{x}_{s-c\epsilon}^\epsilon)] k_j \epsilon(y_s^\epsilon)_j f_p(\mathbf{x}_{s-c_p\epsilon}^\epsilon) ds \\
&\quad + \int_0^t \sum_{p,j,\ell} \frac{\partial}{\partial x_p} [\sigma_{ij}(\mathbf{x}_{s-c\epsilon}^\epsilon)] k_j \epsilon(y_s^\epsilon)_j \sigma_{p\ell}(\mathbf{x}_{s-c_p\epsilon-c\epsilon}^\epsilon) (y_{s-c_p\epsilon}^\epsilon)_\ell ds \\
&\quad - \sum_j \sigma_{ij}(\mathbf{x}_{t-c\epsilon}^\epsilon) k_j \epsilon(y_t^\epsilon)_j + \sum_j \sigma_{ij}(\mathbf{x}_{-c\epsilon}^\epsilon) k_j \epsilon(y_0^\epsilon)_j
\end{aligned} \tag{14}$$

where  $\mathbf{x}_{s-c_p\epsilon}^\epsilon = ((x_{s-c_p\epsilon}^\epsilon)_1, \dots, (x_{s-c_p\epsilon}^\epsilon)_i, \dots, (x_{s-c_p\epsilon}^\epsilon)_m)^\top$  and  $\mathbf{x}_{s-c_p\epsilon-c\epsilon}^\epsilon = ((x_{s-c_p\epsilon-c_1\epsilon}^\epsilon)_1, \dots, (x_{s-c_p\epsilon-c_i\epsilon}^\epsilon)_i, \dots, (x_{s-c_p\epsilon-c_m\epsilon}^\epsilon)_m)^\top$ .

In the above equation, we will show that the boundary terms and the second Lebesgue integral (the one whose integrand contains the factor  $f_p(\mathbf{x}_{s-c_p\epsilon}^\epsilon)$ ) go to zero because  $\epsilon \mathbf{y}^\epsilon$  goes to zero. To prove convergence, we add and subtract like terms without delays on the right hand side of the above equation. In addition, we add and subtract the integral of  $S_i(\mathbf{x})$ , where  $S_i(\mathbf{x})$  is defined in equation (4). In the resulting equation for  $\mathbf{x}^\epsilon$ , we keep with the notation of Lemma 1 by collecting the terms which we will show directly go to zero into a new process called  $\mathbf{U}^\epsilon$ . To this end, let  $\mathbf{U}^\epsilon$  be defined componentwise as

$$\begin{aligned}
(U_t^\epsilon)_i &= - \sum_j \sigma_{ij}(\mathbf{x}_{t-c\epsilon}^\epsilon) k_j \epsilon(y_t^\epsilon)_j + \sum_j \sigma_{ij}(\mathbf{x}_{-c\epsilon}^\epsilon) k_j \epsilon(y_0^\epsilon)_j \\
&\quad + \left( \int_0^t (\boldsymbol{\sigma}(\mathbf{x}_{s-c\epsilon}^\epsilon) - \boldsymbol{\sigma}(\mathbf{x}_s^\epsilon)) d\mathbf{W}_s \right)_i \\
&\quad + \int_0^t \sum_{p,j} \frac{\partial}{\partial x_p} [\sigma_{ij}(\mathbf{x}_{s-c\epsilon}^\epsilon)] k_j \epsilon(y_s^\epsilon)_j f_p(\mathbf{x}_{s-c_p\epsilon}^\epsilon) ds \\
&\quad + \int_0^t \sum_{p,j,\ell} \left[ \frac{\partial}{\partial x_p} [\sigma_{ij}(\mathbf{x}_{s-c\epsilon}^\epsilon)] \sigma_{p\ell}(\mathbf{x}_{s-c_p\epsilon-c\epsilon}^\epsilon) \right. \\
&\quad \quad \left. - \frac{\partial}{\partial x_p} [\sigma_{ij}(\mathbf{x}_s^\epsilon)] \sigma_{p\ell}(\mathbf{x}_s^\epsilon) \right] k_j \epsilon(y_s^\epsilon)_j (y_{s-c_p\epsilon}^\epsilon)_\ell ds
\end{aligned} \tag{15}$$

Then  $\mathbf{x}_t^\epsilon$  can be written componentwise as

$$\begin{aligned}
(x_t^\epsilon)_i &= (x_0)_i + (U_t^\epsilon)_i + \int_0^t f_i(\mathbf{x}_s^\epsilon) ds + \left( \int_0^t \boldsymbol{\sigma}(\mathbf{x}_s^\epsilon) d\mathbf{W}_s \right)_i \\
&+ \int_0^t \sum_{p,j} \frac{\partial}{\partial x_p} [\sigma_{ij}(\mathbf{x}_s^\epsilon)] \sigma_{pj}(\mathbf{x}_s^\epsilon) \frac{1}{2} e^{-\frac{c_p}{k_j}} ds \\
&+ \int_0^t \sum_{p,j} \frac{\partial}{\partial x_p} [\sigma_{ij}(\mathbf{x}_s^\epsilon)] \sigma_{pj}(\mathbf{x}_s^\epsilon) \left[ k_j \epsilon(y_s^\epsilon)_j (y_{s-c_p\epsilon}^\epsilon)_j - \frac{1}{2} e^{-\frac{c_p}{k_j}} \right] ds \\
&+ \int_0^t \sum_{\substack{p,j,\ell \\ j \neq \ell}} \frac{\partial}{\partial x_p} [\sigma_{ij}(\mathbf{x}_s^\epsilon)] \sigma_{p\ell}(\mathbf{x}_s^\epsilon) k_j \epsilon(y_s^\epsilon)_j (y_{s-c_p\epsilon}^\epsilon)_\ell ds.
\end{aligned} \tag{16}$$

We now write equation (16) in the form (9) of Lemma 1 by letting  $\mathbf{h} : \mathbb{R}^m \rightarrow \mathbb{R}^{m \times (1+n+1+mn^2)}$  be the matrix-valued function given by

$$\begin{aligned}
\mathbf{h}(\mathbf{X}) &= \left( \mathbf{f}(\mathbf{X}), \boldsymbol{\sigma}(\mathbf{X}), \mathbf{S}(\mathbf{X}), \boldsymbol{\Lambda}^{11}(\mathbf{X}), \dots, \boldsymbol{\Lambda}^{1n}(\mathbf{X}), \right. \\
&\quad \left. \boldsymbol{\Lambda}^{21}(\mathbf{X}), \dots, \boldsymbol{\Lambda}^{2n}(\mathbf{X}), \boldsymbol{\Lambda}^{n1}(\mathbf{X}), \dots, \boldsymbol{\Lambda}^{nn}(\mathbf{X}) \right)
\end{aligned} \tag{17}$$

where  $\mathbf{S} : \mathbb{R}^m \rightarrow \mathbb{R}^m$  is the vector-valued function defined in equation (4) and  $\boldsymbol{\Lambda}^{j\ell} : \mathbb{R}^m \rightarrow \mathbb{R}^{m \times m}$  is defined componentwise as

$$\Lambda_{ip}^{j\ell}(\mathbf{X}) = \frac{\partial}{\partial X_p} [\sigma_{ij}(\mathbf{X})] \sigma_{p\ell}(\mathbf{X})$$

and by letting  $\mathbf{H}^\epsilon$  be the process, with paths in  $C([0, T], \mathbb{R}^{(1+n+1+mn^2)})$ , given by

$$\mathbf{H}_t^\epsilon = \begin{bmatrix} t \\ \mathbf{W}_t \\ t \\ \mathbf{G}_t^{11} \\ \vdots \\ \mathbf{G}_t^{nn} \end{bmatrix}, \tag{18}$$

where  $\mathbf{G}^{j\ell}$  is the process, with paths in  $C([0, T], \mathbb{R}^m)$ , given by

$$(G_t^{j\ell})_p = \int_0^t \left( k_j \epsilon(y_s^\epsilon)_j (y_{s-c_p\epsilon}^\epsilon)_\ell - E \left[ k_j \epsilon(y_s^\epsilon)_j (y_{s-c_p\epsilon}^\epsilon)_\ell \right] \right) ds. \tag{19}$$

Note that the expectation in the integrand is zero for  $j \neq \ell$  and  $\frac{1}{2} e^{-\frac{c_p}{k_j}}$  for  $j = \ell$  (see proof of Lemma 3).

We will show that the processes  $\mathbf{G}^{j\ell}$  converge to zero in  $L^2$  with respect to  $C([0, T], \mathbb{R}^m)$  for all  $j, \ell = 1, \dots, n$  (Lemma 5). Thus, the limiting process  $\mathbf{H}$  is given by

$$\mathbf{H}_t = \begin{bmatrix} t \\ \mathbf{W}_t \\ t \\ \mathbf{0} \\ \vdots \\ \mathbf{0} \end{bmatrix}. \quad (20)$$

We show in the next subsection that  $\mathbf{U}^\epsilon$ ,  $\mathbf{H}^\epsilon$ , and  $\mathbf{H}$  satisfy the assumptions of Lemma 1. Thus, given the definitions of  $\mathbf{h}$  and  $\mathbf{H}$  above, by computing the right-hand side of

$$\mathbf{x}_t = \mathbf{x}_0 + \int_0^t \mathbf{h}(\mathbf{x}_s) d\mathbf{H}_s \quad (21)$$

we get the limiting SDE (3).

### 3.3 Verifying Conditions of Lemma 1

In this subsection, we complete the proof of Theorem 1 by showing that the conditions of Lemma 1 are satisfied. That is, we show  $\mathbf{U}^\epsilon \rightarrow \mathbf{0}$  in probability with respect to  $C([0, T], \mathbb{R}^m)$ ,  $\mathbf{H}^\epsilon \rightarrow \mathbf{H}$  in probability with respect to  $C([0, T], \mathbb{R}^{(1+n+1+mn^2)})$ , and  $\mathbf{H}^\epsilon$  satisfy Condition 1. We begin with lemmas that we will need. First we show that  $\epsilon \mathbf{y}^\epsilon \rightarrow \mathbf{0}$  in  $L^2$  with respect to  $C([0, T], \mathbb{R}^n)$ . Nelson showed a similar result, namely that  $\sup_{0 \leq t \leq T} \|\epsilon \mathbf{y}_t^\epsilon\| \rightarrow 0$  as  $\epsilon \rightarrow 0$  with probability one [11].

**Lemma 2.** *For each  $\epsilon > 0$ , let  $\mathbf{y}^\epsilon$  be the solution to equation (2) with  $\mathbf{y}_0^\epsilon$  distributed according to the stationary distribution corresponding to (2). There exists  $C' > 0$  independent of  $\epsilon$  such that*

$$E \left[ \left( \sup_{0 \leq t \leq T} \left| \int_0^t e^{-\frac{(t-s)}{k_j \epsilon}} d(W_s)_j \right| \right)^2 \right] \leq C' \epsilon^{1/2}, \quad (22)$$

and this implies that there exists  $C > 0$  independent of  $\epsilon$  such that

$$E \left[ \sup_{0 \leq t \leq T} \|\epsilon \mathbf{y}_t^\epsilon\|^2 \right] \leq C \epsilon^{1/2}. \quad (23)$$

Thus, as  $\epsilon \rightarrow 0$ ,  $\epsilon \mathbf{y}^\epsilon \rightarrow \mathbf{0}$  in  $L^2$ , and therefore in probability, with respect to  $C([0, T], \mathbb{R}^n)$ .

*Proof.* We begin by proving (22) by following the first part of the argument of [13, Lemma 3.7]. We fix  $\alpha \in (0, \frac{1}{2})$  and use the factorization method from [1] (see also [2, Section 5.3]) to rewrite

$$\begin{aligned} I(t) &= \int_0^t e^{-\frac{(t-s)}{k_j \epsilon}} d(W_s)_j \\ &= \frac{\sin(\pi\alpha)}{\pi} \int_0^t e^{-\frac{(t-s)}{k_j \epsilon}} (t-s)^{\alpha-1} Y(s) ds, \end{aligned}$$

where

$$Y(s) = \int_0^s e^{-\frac{(s-u)}{k_j \epsilon}} (s-u)^{-\alpha} d(W_u)_j .$$

and we used the identity

$$\int_u^t (t-s)^{\alpha-1} (s-u)^{-\alpha} ds = \frac{\pi}{\sin(\pi\alpha)} , \quad 0 < \alpha < 1 .$$

We fix  $m > \frac{1}{2\alpha}$  and use the Hölder inequality:

$$|I(t)|^{2m} \leq C_1 \left( \int_0^t \left| e^{-\frac{(t-s)}{k_j \epsilon}} (t-s)^{\alpha-1} \right|^{\frac{2m}{2m-1}} ds \right)^{2m-1} \int_0^t |Y(s)|^{2m} ds .$$

Using the change of variables  $z = \frac{2m}{2m-1} \frac{(t-s)}{k_j \epsilon}$ , we have

$$\begin{aligned} & \int_0^t e^{-\frac{2m}{2m-1} \frac{(t-s)}{k_j \epsilon}} (t-s)^{\frac{2m}{2m-1}(\alpha-1)} ds \\ &= \left( \frac{k_j(2m-1)}{2m} \right)^{\frac{2m\alpha-1}{2m-1}} \epsilon^{\frac{2m\alpha-1}{2m-1}} \int_0^{\frac{t}{k_j \epsilon} \frac{2m}{2m-1}} e^{-z} z^{\frac{2m}{2m-1}(\alpha-1)} dz \\ &\leq C_2 \epsilon^{\frac{2m\alpha-1}{2m-1}} \int_0^\infty e^{-z} z^{\frac{2m}{2m-1}(\alpha-1)} dz \\ &\leq C_3 \epsilon^{\frac{2m\alpha-1}{2m-1}} \end{aligned}$$

where in the above we have used the fact that  $e^{-z} z^{\frac{2m}{2m-1}(\alpha-1)} \in L^1(\mathbb{R}^+)$  since  $m > \frac{1}{2\alpha}$ . Therefore, we have

$$E \left[ \sup_{0 \leq t \leq T} |I(t)|^{2m} \right] \leq C_4 \epsilon^{2m\alpha-1} E \left[ \int_0^T |Y(s)|^{2m} ds \right] .$$

Then, letting  $\frac{1}{4} < \alpha < \frac{1}{2}$  and  $m = 2$ , we have

$$\begin{aligned} E[|Y(t)|^4] &= 3 \left( E \left[ (Y(t))^2 \right] \right)^2 \quad \text{since } Y(t) \text{ is a zero-mean Gaussian} \\ &= 3 \left( \int_0^t e^{-\frac{2(t-u)}{k_j \epsilon}} (t-u)^{-2\alpha} du \right)^2 \quad \text{by the Itô isometry} \\ &= 3 \left( \frac{k_j \epsilon}{2} \right)^{2(1-2\alpha)} \left( \int_0^{\frac{2t}{k_j \epsilon}} e^{-s} s^{-2\alpha} ds \right)^2 \\ &\leq C_5 \epsilon^{2(1-2\alpha)} \quad \text{using the fact that } e^{-s} s^{-2\alpha} \in L^1(\mathbb{R}^+) \text{ since } \alpha < \frac{1}{2} . \end{aligned}$$

Thus,

$$E \left[ \sup_{0 \leq t \leq T} |I(t)|^4 \right] \leq C_6 \epsilon ,$$

where  $C_6$  is a constant that depends on  $T$ , so we get (22) by the Cauchy-Schwarz inequality:

$$E \left[ \sup_{0 \leq t \leq T} |I(t)|^2 \right] \leq \left( E \left[ \sup_{0 \leq t \leq T} |I(t)|^4 \right] \right)^{1/2} \leq C' \epsilon^{1/2} .$$

Now we prove (23). The solution of (2) is

$$(y_t^\epsilon)_j = e^{-\frac{t}{k_j \epsilon}} (y_0^\epsilon)_j + \frac{1}{k_j \epsilon} \int_0^t e^{-\frac{(t-s)}{k_j \epsilon}} d(W_s)_j .$$

Thus, we have

$$E \left[ \sup_{0 \leq t \leq T} |(\epsilon y_t^\epsilon)_j|^2 \right] \leq 2E \left[ \sup_{0 \leq t \leq T} \left| e^{-\frac{t}{k_j \epsilon}} (\epsilon y_0^\epsilon)_j \right|^2 \right] + \frac{2}{k_j^2} E \left[ \sup_{0 \leq t \leq T} \left( \int_0^t e^{-\frac{(t-s)}{k_j \epsilon}} d(W_s)_j \right)^2 \right] .$$

For the first term we use that  $(y_0^\epsilon)_j$  is distributed according to the stationary distribution corresponding to (2) and thus has mean zero and variance  $E[|(y_0^\epsilon)_j|^2] = (2k_j \epsilon)^{-1}$ :

$$E \left[ \sup_{0 \leq t \leq T} \left| e^{-\frac{t}{k_j \epsilon}} (\epsilon y_0^\epsilon)_j \right|^2 \right] \leq E \left[ |(\epsilon y_0^\epsilon)_j|^2 \right] = \frac{\epsilon}{2k_j}$$

The bound (23) then follows from this bound together with (22).  $\square$

The next lemma is elementary but we include its proof for completeness.

**Lemma 3.** *For each  $\epsilon > 0$ , let  $\mathbf{y}^\epsilon$  be defined on  $t \geq 0$  as the stationary solution to equation (2) with  $\mathbf{y}_0^\epsilon$  distributed according to the stationary distribution corresponding to (2), and let  $\mathbf{y}^\epsilon$  be given on  $[t_-, 0]$  by  $(\mathbf{y}^-)^\epsilon$  as defined in the statement of Theorem 1 (so that  $(\mathbf{y}^-)^\epsilon_0 = \mathbf{y}_0^\epsilon$ ). Then there exists  $C > 0$  such that for all  $\epsilon > 0$ ,*

$$E \left[ |\epsilon(y_t^\epsilon)_j (y_s^\epsilon)_\ell|^2 \right] < C$$

for all  $1 \leq j, \ell \leq n$  and  $t, s \geq t_-$

*Proof.* Since  $(y_0^\epsilon)_j$  is distributed according to the stationary distribution corresponding to (2),  $(y_t^\epsilon)_j$  is a mean zero Gaussian random variable with variance

$$E \left[ |(y_t^\epsilon)_j|^2 \right] = \frac{1}{2k_j \epsilon} . \quad (24)$$

Furthermore,  $((y_t^\epsilon)_j, (y_s^\epsilon)_j)$  is jointly Gaussian with covariance

$$E[(y_t^\epsilon)_j (y_s^\epsilon)_j] = \frac{1}{2k_j \epsilon} e^{-\frac{|t-s|}{k_j \epsilon}} . \quad (25)$$

Recall that for  $j \neq \ell$ ,  $(y_t^\epsilon)_j$  is independent of  $(y_s^\epsilon)_\ell$  for all  $t, s \geq t_-$ . Let  $\delta_{j\ell}$  denote the Kronecker delta. Then using Wick's Theorem for the centered, Gaussian random variables  $(y_t^\epsilon)_j$  and  $(y_s^\epsilon)_\ell$  [7, Theorem 1.28], along with (24) and (25), we have

$$\begin{aligned} E \left[ |\epsilon(y_t^\epsilon)_j (y_s^\epsilon)_\ell|^2 \right] &= E \left[ |\sqrt{\epsilon}(y_t^\epsilon)_j|^2 \right] E \left[ |\sqrt{\epsilon}(y_s^\epsilon)_\ell|^2 \right] + 2 \left( E \left[ \epsilon(y_t^\epsilon)_j (y_s^\epsilon)_\ell \right] \right)^2 \\ &= \frac{1}{4k_j k_\ell} + \frac{\delta_{j\ell}}{2k_j^2} e^{-\frac{2|t-s|}{k_j \epsilon}} \leq C \end{aligned}$$

where  $C$  is a constant independent of  $\epsilon$ . □

**Lemma 4.** *For each  $\epsilon > 0$ , let  $\mathbf{x}^\epsilon$  be defined as in the statement of Theorem 1 where  $\mathbf{f}$  is bounded and  $\boldsymbol{\sigma}$  is bounded with bounded derivatives. Also, for  $\mathbf{t} = (t_1, \dots, t_m) \in [0, T]^m$  and  $\mathbf{h} = (h_1, \dots, h_m) \in [0, 1]^m$ , let  $\mathbf{x}_\mathbf{t}^\epsilon = ((x_{t_1}^\epsilon)_1, \dots, (x_{t_m}^\epsilon)_m)$ , let  $\mathbf{x}_{\mathbf{t}+\mathbf{h}}^\epsilon = ((x_{t_1+h_1}^\epsilon)_1, \dots, (x_{t_m+h_m}^\epsilon)_m)$ , and let  $h_M = \max_{1 \leq i \leq m} h_i$ . Then there exists  $C$  independent of  $\epsilon$  such that for all  $\mathbf{t} \in [0, T]^m$ ,  $\mathbf{h} \in [0, 1]^m$ , and  $\epsilon > 0$ ,*

$$E \left[ \|\mathbf{x}_{\mathbf{t}+\mathbf{h}}^\epsilon - \mathbf{x}_\mathbf{t}^\epsilon\|^2 \right] \leq C \left( h_M + \epsilon^{1/2} \right)$$

*Proof.* From equation (14) and the Cauchy-Schwarz inequality, we have

$$\begin{aligned} E \left[ |(x_{t_i+h_i}^\epsilon)_i - (x_{t_i}^\epsilon)_i|^2 \right] &\leq \\ 6E \left[ \left( \int_{t_i}^{t_i+h_i} f_i(\mathbf{x}_s^\epsilon) ds \right)^2 \right] &+ 6E \left[ \left( \left( \int_{t_i}^{t_i+h_i} \boldsymbol{\sigma}(\mathbf{x}_{s-\mathbf{c}\epsilon}) d\mathbf{W}_s \right)_i \right)^2 \right] \\ + 6E \left[ \left( \int_{t_i}^{t_i+h_i} \sum_{p,j} \frac{\partial}{\partial x_p} [\sigma_{ij}(\mathbf{x}_{s-\mathbf{c}\epsilon})] k_j \epsilon (y_s^\epsilon)_j f_p(\mathbf{x}_{s-\mathbf{c}_p\epsilon}) ds \right)^2 \right] \\ + 6E \left[ \left( \int_{t_i}^{t_i+h_i} \sum_{p,j,\ell} \frac{\partial}{\partial x_p} [\sigma_{ij}(\mathbf{x}_{s-\mathbf{c}\epsilon})] k_j \epsilon (y_s^\epsilon)_j \sigma_{p\ell}(\mathbf{x}_{s-\mathbf{c}_p\epsilon-\mathbf{c}\epsilon}) (y_{s-\mathbf{c}_p\epsilon}^\epsilon)_\ell ds \right)^2 \right] \\ + 6E \left[ \left( \sum_j \sigma_{ij}(\mathbf{x}_{t_i+h_i-\mathbf{c}\epsilon}) k_j \epsilon (y_{t_i+h_i}^\epsilon)_j \right)^2 \right] &+ 6E \left[ \left( \sum_j \sigma_{ij}(\mathbf{x}_{t_i-\mathbf{c}\epsilon}) k_j \epsilon (y_{t_i}^\epsilon)_j \right)^2 \right] \end{aligned}$$

Using the boundedness of  $f_i$ , the Itô isometry, and the Cauchy-Schwarz inequality,

ity then gives

$$\begin{aligned}
& E \left[ |(x_{t_i+h_i}^\epsilon)_i - (x_{t_i}^\epsilon)_i|^2 \right] \leq \\
& C_1 h_i^2 + 6C_2 \int_{t_i}^{t_i+h_i} E \left[ \|\sigma(\mathbf{x}_{s-c\epsilon}^\epsilon)\|^2 \right] ds \\
& + 6h_i E \left[ \int_{t_i}^{t_i+h_i} \left( \sum_{p,j} \frac{\partial}{\partial x_p} [\sigma_{ij}(\mathbf{x}_{s-c\epsilon}^\epsilon)] k_j \epsilon(y_s^\epsilon)_j f_p(\mathbf{x}_{s-c_p\epsilon}^\epsilon) \right)^2 ds \right] \\
& + 6h_i E \left[ \int_{t_i}^{t_i+h_i} \left( \sum_{p,j,\ell} \frac{\partial}{\partial x_p} [\sigma_{ij}(\mathbf{x}_{s-c\epsilon}^\epsilon)] k_j \epsilon(y_s^\epsilon)_j \sigma_{p\ell}(\mathbf{x}_{s-c_p\epsilon-c\epsilon}^\epsilon) (y_{s-c_p\epsilon}^\epsilon)_\ell \right)^2 ds \right] \\
& + 6n E \left[ \sum_j \left( \sigma_{ij}(\mathbf{x}_{t_i+h_i-c\epsilon}^\epsilon) k_j \epsilon(y_{t_i+h_i}^\epsilon)_j \right)^2 \right] + 6n E \left[ \sum_j \left( \sigma_{ij}(\mathbf{x}_{t_i-c\epsilon}^\epsilon) k_j \epsilon(y_{t_i}^\epsilon)_j \right)^2 \right]
\end{aligned}$$

Using the boundedness of  $\mathbf{f}$  and the boundedness of  $\sigma$  and its derivatives, we get

$$\begin{aligned}
E \left[ |(x_{t_i+h_i}^\epsilon)_i - (x_{t_i}^\epsilon)_i|^2 \right] & \leq C_1 h_i^2 + C_3 h_i + 6C_4 h_i \int_{t_i}^{t_i+h_i} E \left[ \sum_j \left( \epsilon(y_s^\epsilon)_j \right)^2 \right] ds \\
& + 6C_5 h_i \int_{t_i}^{t_i+h_i} E \left[ \sum_{j,\ell} \left( \epsilon(y_s^\epsilon)_j (y_{s-c_p\epsilon}^\epsilon)_\ell \right)^2 \right] ds \\
& + 6nC_6 E \left[ \sum_j \left( \epsilon(y_{t_i+h_i}^\epsilon)_j \right)^2 \right] + 6nC_6 E \left[ \sum_j \left( \epsilon(y_{t_i}^\epsilon)_j \right)^2 \right]
\end{aligned}$$

Thus, using Lemma 2 and Lemma 3,

$$E \left[ |(x_{t_i+h_i}^\epsilon)_i - (x_{t_i}^\epsilon)_i|^2 \right] \leq C_1 h_i^2 + C_3 h_i + 6C_7 \epsilon^{1/2} h_i^2 + 6C_8 h_i^2 + C_9 \epsilon^{1/2}$$

from which the statement follows.  $\square$

### 3.3.1 $U^\epsilon$ converges to zero

Now we are ready to prove that  $U^\epsilon$  goes to zero in probability with respect to  $C([0, T], \mathbb{R}^m)$  as  $\epsilon \rightarrow 0$ . To do this, we split  $(U^\epsilon)_i$  into two types of terms.

Recall that  $(U^\epsilon)_i$  is defined in equation (15) as

$$\begin{aligned}
(U_t^\epsilon)_i &= \underbrace{\left( \int_0^t (\sigma(\mathbf{x}_{s-c\epsilon}^\epsilon) - \sigma(\mathbf{x}_s^\epsilon)) d\mathbf{W}_s \right)_i}_I \\
&\quad + \underbrace{\int_0^t \sum_{p,j,\ell} \left[ \frac{\partial}{\partial x_p} [\sigma_{ij}(\mathbf{x}_{s-c\epsilon}^\epsilon)] \sigma_{p\ell}(\mathbf{x}_{s-c_p\epsilon-c\epsilon}^\epsilon) - \frac{\partial}{\partial x_p} [\sigma_{ij}(\mathbf{x}_s^\epsilon)] \sigma_{p\ell}(\mathbf{x}_s^\epsilon) \right] k_j \epsilon(y_s^\epsilon)_j (y_{s-c_p\epsilon}^\epsilon)_\ell ds}_I \\
&\quad - \underbrace{\sum_j \sigma_{ij}(\mathbf{x}_{t-c\epsilon}^\epsilon) k_j \epsilon(y_t^\epsilon)_j + \sum_j \sigma_{ij}(\mathbf{x}_{-c\epsilon}^\epsilon) k_j \epsilon(y_0^\epsilon)_j}_{II} \\
&\quad + \underbrace{\int_0^t \sum_{p,j} \frac{\partial}{\partial x_p} [\sigma_{ij}(\mathbf{x}_{s-c\epsilon}^\epsilon)] k_j \epsilon(y_s^\epsilon)_j f_p(\mathbf{x}_{s-c_p\epsilon}^\epsilon) ds}_{II}
\end{aligned}$$

We will prove that these terms all converge to zero as  $\epsilon \rightarrow 0$ . Convergence of the type  $I$  terms to zero will be a consequence of Lemma 4 and the Lipschitz continuity of  $\sigma$  and its first derivatives. The proof of convergence of type  $II$  terms to zero will follow from  $\epsilon \mathbf{y}^\epsilon \rightarrow \mathbf{0}$  as  $\epsilon \rightarrow 0$ .

We start with the first two terms of  $(U_t^\epsilon)_i$ , i.e., the type  $I$  terms. Recall that for  $p \geq 1$ , convergence in  $L^p$  with respect to  $C([0, T], \mathbb{R})$  implies convergence in probability with respect to  $C([0, T], \mathbb{R})$ . Thus, to show that these terms converge to zero in probability with respect to  $C([0, T], \mathbb{R})$ , we show that the first term goes to zero in  $L^2$  with respect to  $C([0, T], \mathbb{R})$  and the second term goes to zero in  $L^1$  with respect to  $C([0, T], \mathbb{R})$ . We note that it is possible to show that the second term goes to zero in  $L^2$  with respect to  $C([0, T], \mathbb{R})$ , but it would take slightly more work. Thus, it suffices to show that for every  $i, j, \ell, p$ ,

$$\lim_{\epsilon \rightarrow 0} E \left[ \left( \sup_{0 \leq t \leq T} \left| \int_0^t (\sigma_{ij}(\mathbf{x}_{s-c\epsilon}^\epsilon) - \sigma_{ij}(\mathbf{x}_s^\epsilon)) d(W_s)_j \right| \right)^2 \right] = 0 \quad (26)$$

$$\begin{aligned}
\text{and} \quad \lim_{\epsilon \rightarrow 0} E \left[ \sup_{0 \leq t \leq T} \left| \int_0^t \left[ \frac{\partial}{\partial x_p} [\sigma_{ij}(\mathbf{x}_{s-c\epsilon}^\epsilon)] \sigma_{p\ell}(\mathbf{x}_{s-c_p\epsilon-c\epsilon}^\epsilon) \right. \right. \right. \\
\left. \left. \left. - \frac{\partial}{\partial x_p} [\sigma_{ij}(\mathbf{x}_s^\epsilon)] \sigma_{p\ell}(\mathbf{x}_s^\epsilon) \right] k_j \epsilon(y_s^\epsilon)_j (y_{s-c_p\epsilon}^\epsilon)_\ell ds \right| \right] = 0. \quad (27)
\end{aligned}$$

First we show (26). We will use Lipschitz continuity of  $\sigma_{ij}$  for all  $i, j$ , which follows from the assumption of Theorem 1 that  $\sigma$  has bounded first derivatives. Also, we note that the Itô integral in (26) is a martingale because the integrand is still nonanticipating in the presence of the delays. Thus, by Doob's maximal



inequality and the Itô isometry, we have

$$\begin{aligned}
& E \left[ \left( \sup_{0 \leq t \leq T} \left| \int_0^t (\sigma_{ij}(\mathbf{x}_{s-c\epsilon}^\epsilon) - \sigma_{ij}(\mathbf{x}_s^\epsilon)) d(W_s)_j \right| \right)^2 \right] \\
& \leq 2E \left[ \left( \int_0^T (\sigma_{ij}(\mathbf{x}_{s-c\epsilon}^\epsilon) - \sigma_{ij}(\mathbf{x}_s^\epsilon)) d(W_s)_j \right)^2 \right] \\
& = 2 \int_0^T E \left[ (\sigma_{ij}(\mathbf{x}_{s-c\epsilon}^\epsilon) - \sigma_{ij}(\mathbf{x}_s^\epsilon))^2 \right] ds.
\end{aligned}$$

We bound the above integral by using the Lipschitz continuity of  $\sigma_{ij}$  and Lemma 4. Since Lemma 4 applies only to differences of values of the process  $\mathbf{x}^\epsilon$  at nonnegative times, we split the above integral into two terms. The first term involves values of  $\mathbf{x}^\epsilon$  at negative times, i.e. values of the past condition  $\mathbf{x}^-$  and the second term only involves values of  $\mathbf{x}^\epsilon$  at nonnegative times. Letting  $c_M = \max_{1 \leq i \leq m} c_i$ , we have

$$\begin{aligned}
& E \left[ \left( \sup_{0 \leq t \leq T} \left| \int_0^t (\sigma_{ij}(\mathbf{x}_{s-c\epsilon}^\epsilon) - \sigma_{ij}(\mathbf{x}_s^\epsilon)) d(W_s)_j \right| \right)^2 \right] \\
& \leq 2 \int_0^{c_M \epsilon} E \left[ (\sigma_{ij}(\mathbf{x}_{s-c\epsilon}^\epsilon) - \sigma_{ij}(\mathbf{x}_s^\epsilon))^2 \right] ds + 2 \int_{c_M \epsilon}^T E \left[ (M \|\mathbf{x}_{s-c\epsilon}^\epsilon - \mathbf{x}_s^\epsilon\|)^2 \right] ds \\
& \quad \text{where } M \text{ is the Lipschitz constant for } \sigma_{ij} \\
& \leq C_1 \epsilon + C_2 T (\epsilon + \epsilon^{1/2})
\end{aligned}$$

using the boundedness of  $\sigma_{ij}$  and Lemma 4, from which (26) follows. To prove the bound (27):

$$\begin{aligned}
& E \left[ \sup_{0 \leq t \leq T} \left| \int_0^t \left[ \frac{\partial}{\partial x_p} [\sigma_{ij}(\mathbf{x}_{s-c\epsilon}^\epsilon)] \sigma_{p\ell}(\mathbf{x}_{s-c_p\epsilon-c\epsilon}^\epsilon) - \frac{\partial}{\partial x_p} [\sigma_{ij}(\mathbf{x}_s^\epsilon)] \sigma_{p\ell}(\mathbf{x}_s^\epsilon) \right] k_j \epsilon(y_s^\epsilon)_j (y_{s-c_p\epsilon}^\epsilon)_\ell ds \right| \right] \\
& \leq \int_0^T E \left[ \left| \left[ \frac{\partial}{\partial x_p} [\sigma_{ij}(\mathbf{x}_{s-c\epsilon}^\epsilon)] \sigma_{p\ell}(\mathbf{x}_{s-c_p\epsilon-c\epsilon}^\epsilon) - \frac{\partial}{\partial x_p} [\sigma_{ij}(\mathbf{x}_s^\epsilon)] \sigma_{p\ell}(\mathbf{x}_s^\epsilon) \right] k_j \epsilon(y_s^\epsilon)_j (y_{s-c_p\epsilon}^\epsilon)_\ell \right| \right] ds \\
& \leq \int_0^T \left( E \left[ \left( \frac{\partial}{\partial x_p} [\sigma_{ij}(\mathbf{x}_{s-c\epsilon}^\epsilon)] \sigma_{p\ell}(\mathbf{x}_{s-c_p\epsilon-c\epsilon}^\epsilon) - \frac{\partial}{\partial x_p} [\sigma_{ij}(\mathbf{x}_s^\epsilon)] \sigma_{p\ell}(\mathbf{x}_s^\epsilon) \right)^2 \right] \right)^{1/2} \\
& \quad \times \left( E \left[ \left( k_j \epsilon(y_s^\epsilon)_j (y_{s-c_p\epsilon}^\epsilon)_\ell \right)^2 \right] \right)^{1/2} ds \\
& \leq C_1 \int_0^T \left( E \left[ \left( \frac{\partial}{\partial x_p} [\sigma_{ij}(\mathbf{x}_{s-c\epsilon}^\epsilon)] \sigma_{p\ell}(\mathbf{x}_{s-c_p\epsilon-c\epsilon}^\epsilon) - \frac{\partial}{\partial x_p} [\sigma_{ij}(\mathbf{x}_s^\epsilon)] \sigma_{p\ell}(\mathbf{x}_s^\epsilon) \right)^2 \right] \right)^{1/2} ds
\end{aligned}$$

by Lemma 3. Again, in order to apply Lemma 4, we split the above integral into a part that involves values of the past condition  $\mathbf{x}^-$  and a part that only

involves values of  $\mathbf{x}^\epsilon$  at nonnegative times. We have

$$\begin{aligned}
& E \left[ \sup_{0 \leq t \leq T} \left| \int_0^t \left[ \frac{\partial}{\partial x_p} [\sigma_{ij}(\mathbf{x}_{s-c\epsilon}^\epsilon)] \sigma_{p\ell}(\mathbf{x}_{s-c_p\epsilon-c\epsilon}^\epsilon) - \frac{\partial}{\partial x_p} [\sigma_{ij}(\mathbf{x}_s^\epsilon)] \sigma_{p\ell}(\mathbf{x}_s^\epsilon) \right] k_j \epsilon (y_s^\epsilon)_j (y_{s-c_p\epsilon}^\epsilon)_\ell ds \right| \right] \\
& \leq C_1 \int_0^{2c_M\epsilon} \left( E \left[ \left( \frac{\partial}{\partial x_p} [\sigma_{ij}(\mathbf{x}_{s-c\epsilon}^\epsilon)] \sigma_{p\ell}(\mathbf{x}_{s-c_p\epsilon-c\epsilon}^\epsilon) - \frac{\partial}{\partial x_p} [\sigma_{ij}(\mathbf{x}_s^\epsilon)] \sigma_{p\ell}(\mathbf{x}_s^\epsilon) \right)^2 \right] \right)^{1/2} ds \\
& + C_1 \int_{2c_M\epsilon}^T \left( E \left[ 2 \left( \frac{\partial}{\partial x_p} [\sigma_{ij}(\mathbf{x}_{s-c\epsilon}^\epsilon)] \sigma_{p\ell}(\mathbf{x}_{s-c_p\epsilon-c\epsilon}^\epsilon) - \frac{\partial}{\partial x_p} [\sigma_{ij}(\mathbf{x}_{s-c\epsilon}^\epsilon)] \sigma_{p\ell}(\mathbf{x}_s^\epsilon) \right)^2 \right. \right. \\
& \quad \left. \left. + 2 \left( \frac{\partial}{\partial x_p} [\sigma_{ij}(\mathbf{x}_{s-c\epsilon}^\epsilon)] \sigma_{p\ell}(\mathbf{x}_s^\epsilon) - \frac{\partial}{\partial x_p} [\sigma_{ij}(\mathbf{x}_s^\epsilon)] \sigma_{p\ell}(\mathbf{x}_s^\epsilon) \right)^2 \right] \right)^{1/2} ds \\
& \leq C_2\epsilon + C_3 \int_{2c\epsilon}^T \left( E \left[ \|\mathbf{x}_{s-c_p\epsilon-c\epsilon}^\epsilon - \mathbf{x}_s^\epsilon\|^2 + \|\mathbf{x}_{s-c\epsilon}^\epsilon - \mathbf{x}_s^\epsilon\|^2 \right] \right)^{1/2} ds \\
& \quad (\text{using the boundedness and Lipschitz continuity of } \boldsymbol{\sigma} \text{ and its first derivatives}) \\
& \leq C_2\epsilon + C_4T(\epsilon + \epsilon^{1/2})^{1/2}
\end{aligned}$$

by Lemma 4, which gives (27).

For the type II terms, it suffices to show that for every  $i, j, p$ ,

$$\lim_{\epsilon \rightarrow 0} E \left[ \sup_{0 \leq t \leq T} |\sigma_{ij}(\mathbf{x}_{t-c\epsilon}^\epsilon) k_j \epsilon (y_t^\epsilon)_j|^2 \right] = 0 \quad (28)$$

$$\lim_{\epsilon \rightarrow 0} E \left[ \sup_{0 \leq t \leq T} |\sigma_{ij}(\mathbf{x}_{-c\epsilon}^\epsilon) k_j \epsilon (y_0^\epsilon)_j|^2 \right] = 0 \quad (29)$$

$$\text{and } \lim_{\epsilon \rightarrow 0} E \left[ \sup_{0 \leq t \leq T} \left| \int_0^t \frac{\partial}{\partial x_p} [\sigma_{ij}(\mathbf{x}_{s-c\epsilon}^\epsilon)] k_j \epsilon (y_s^\epsilon)_j f_p(\mathbf{x}_{s-c_p\epsilon}^\epsilon) ds \right|^2 \right] = 0. \quad (30)$$

Equations (28) and (29) follow immediately from the boundedness of  $\sigma_{ij}$  and Lemma 2. To prove equation (30), we first use the boundedness of  $\frac{\partial}{\partial x_p} \sigma_{ij}$  and  $f_p$  and then apply the Cauchy-Schwarz inequality:

$$\begin{aligned}
& E \left[ \sup_{0 \leq t \leq T} \left| \int_0^t \frac{\partial}{\partial x_p} [\sigma_{ij}(\mathbf{x}_{s-c\epsilon}^\epsilon)] k_j \epsilon (y_s^\epsilon)_j f_p(\mathbf{x}_{s-c_p\epsilon}^\epsilon) ds \right|^2 \right] \\
& \leq CE \left[ \left( \int_0^T |\epsilon (y_s^\epsilon)_j| ds \right)^2 \right] \\
& \leq CT \int_0^T E \left[ |\epsilon (y_s^\epsilon)_j|^2 \right] ds
\end{aligned}$$

which goes to zero by Lemma 2.

Therefore  $\mathbf{U}^\epsilon \rightarrow \mathbf{0}$  as  $\epsilon \rightarrow 0$  in probability with respect to  $C([0, T], \mathbb{R}^m)$ , as claimed

□

### 3.3.2 $H^\epsilon \rightarrow H$

Here we show that  $H^\epsilon \rightarrow H$  in  $L^2$ , and thus in probability, with respect to  $C([0, T], \mathbb{R}^{(1+n+1+mn^2)})$ . Note that the first three components of  $H_t^\epsilon$ , defined in (18), are independent of  $\epsilon$ . Thus, it suffices to show that  $(G^{j\ell})_p \rightarrow 0$  in  $L^2$  with respect to  $C([0, T], \mathbb{R})$  for all  $j, \ell = 1, \dots, n$  and  $p = 1, \dots, m$ .

A heuristic argument that provides some insight into why  $(G^{j\ell})_p$  converges to zero is as follows. Recall that  $(G_t^{j\ell})_p$  is defined as

$$(G_t^{j\ell})_p = \int_0^t \left( k_j \epsilon (y_s^\epsilon)_j (y_{s-c_p\epsilon}^\epsilon)_\ell - E \left[ k_j \epsilon (y_s^\epsilon)_j (y_{s-c_p\epsilon}^\epsilon)_\ell \right] \right) ds. \quad (31)$$

For each  $j$ , define the new process  $(z)_j$  by  $(z_r)_j = \sqrt{\epsilon} (y_{\epsilon r}^\epsilon)_j$ . Then  $(z)_j$  solves the  $\epsilon$ -independent equation

$$d(z_r)_j = -\frac{1}{k_j} (z_r)_j dr + \frac{1}{k_j} d(\tilde{W}_r)_j.$$

with the Wiener process  $\tilde{W}_r = \epsilon^{-\frac{1}{2}} W_{\epsilon r}$ . In terms of the process  $(z)_j$ ,  $(G_t^{j\ell})_p$  can be written as

$$(G_t^{j\ell})_p = \epsilon \int_0^{t/\epsilon} \left( k_j (z_u)_j (z_{u-c_p})_\ell - E \left[ k_j (z_u)_j (z_{u-c_p})_\ell \right] \right) du.$$

The above integral can be thought of as the sum of  $m = O(1/\epsilon)$  identically distributed random variables with zero mean. Furthermore, these random variables are weakly correlated, since the covariance function for  $(z)_j$  decays exponentially with an exponential decay constant of order 1. Thus, we expect the  $L^2$ -norm of this sum to grow about as fast as  $O(1/\epsilon)$ . Since  $(G_t^{j\ell})_p$  is equal to this integral multiplied by  $\epsilon$ , we expect  $(G_t^{j\ell})_p$  to converge to zero as  $\epsilon \rightarrow 0$  for all  $t \in [0, T]$ . For fixed  $t \in [0, T]$ , convergence of  $(G_t^{j\ell})_p$  to zero in  $L^2$  can be shown by expressing the square of the integral as a double integral and then using Wick's theorem to compute the expectation of the terms in the integrand. However, to control the supremum norm, more work is required, and this is the content of the next lemma.

**Lemma 5.** *For each  $\epsilon > 0$ , let  $\mathbf{y}^\epsilon$  be defined on  $t \geq 0$  as the solution to equation (2) with  $\mathbf{y}_0^\epsilon$  distributed according to the stationary distribution corresponding to (2), and let  $\mathbf{y}^\epsilon$  be given on  $[t_-, 0]$  by  $(\mathbf{y}^-)^\epsilon$  as defined in the statement of Theorem 1. Then for all  $j, \ell, p$ ,*

$$\lim_{\epsilon \rightarrow 0} E \left[ \left( \sup_{0 \leq t \leq T} \left| \int_0^t \left( k_j \epsilon (y_u^\epsilon)_j (y_{u-c_p\epsilon}^\epsilon)_\ell - \frac{1}{2} e^{-\frac{c_p}{k_j}} \right) du \right| \right)^2 \right] = 0 \quad (32)$$

and

$$\lim_{\epsilon \rightarrow 0} E \left[ \left( \sup_{0 \leq t \leq T} \left| \int_0^t k_j \epsilon (y_u^\epsilon)_j (y_{u-c_p\epsilon}^\epsilon)_\ell du \right| \right)^2 \right] = 0 \quad \text{for } j \neq \ell \quad (33)$$

To show this, a Riemann sum approximation is used in order to apply the following maximal inequality for sums of a stationary sequence of random variables.

**Lemma 6.** [15, Proposition 2.3] *Let  $\{X_i : i \in \mathbb{N}\}$  be a stationary sequence of random variables and let  $S_n = \sum_{i=1}^n X_i$ . Let  $n \in \mathbb{N}$  and let  $r = \lceil \log_2(n) \rceil$  (i.e.  $r$  is the smallest integer greater than or equal to  $\log_2(n)$ ). Then*

$$E \left[ \max_{1 \leq i \leq n} S_i^2 \right] \leq n \left( 2\sqrt{E[X_1^2]} + (1 + \sqrt{2}) \Delta_r \right)^2$$

where

$$\Delta_r = \sum_{j=0}^{r-1} \left| \frac{E[S_{2^j}]}{2^{j/2}} \right|$$

*Proof of Lemma 5.* Let

$$\Psi_{j\ell p} = E \left[ k_j \epsilon(y_u^\epsilon)_j (y_{u-c_p\epsilon}^\epsilon)_\ell \right] = \frac{\delta_{j\ell}}{2} e^{-\frac{c_p}{k_j}}.$$

Then showing both (32) and (33) is equivalent to showing that for all  $j, \ell, p$ ,

$$\lim_{\epsilon \rightarrow 0} E \left[ \left( \sup_{0 \leq t \leq T} \left| \int_0^t \left( k_j \epsilon(y_u^\epsilon)_j (y_{u-c_p\epsilon}^\epsilon)_\ell - \Psi_{j\ell p} \right) du \right| \right)^2 \right] = 0. \quad (34)$$

For each  $j$ , define the new process  $(z)_j$  by  $(z_r)_j = \sqrt{\epsilon}(y_{\epsilon r}^\epsilon)_j$ . Then  $(z)_j$  solves the  $\epsilon$ -independent equation

$$d(z_r)_j = -\frac{1}{k_j}(z_r)_j dr + \frac{1}{k_j} d(\tilde{W}_r)_j, \quad (z_0)_j = \sqrt{\epsilon}(y_0^\epsilon)_j$$

with the Wiener process  $\tilde{W}_r = \epsilon^{-\frac{1}{2}} W_{\epsilon r}$ . Letting  $r = \frac{u}{\epsilon}$  and  $N = \frac{1}{\epsilon}$  gives

$$\begin{aligned} \int_0^t \left( k_j \epsilon(y_u^\epsilon)_j (y_{u-c_p\epsilon}^\epsilon)_\ell - \Psi_{j\ell p} \right) du &= \epsilon \int_0^{t/\epsilon} \left( k_j \epsilon(y_{\epsilon r}^\epsilon)_j (y_{\epsilon r-c_p\epsilon}^\epsilon)_\ell - \Psi_{j\ell p} \right) dr \\ &= \epsilon \int_0^{t/\epsilon} \left( k_j (z_r)_j (z_{r-c_p})_\ell - \Psi_{j\ell p} \right) dr \\ &= \frac{1}{N} \int_0^{Nt} \left( k_j (z_r)_j (z_{r-c_p})_\ell - \Psi_{j\ell p} \right) dr \end{aligned}$$

We approximate the above integral by a Riemann sum. Let  $\{t_i\}_{i=1, \dots, m}$  be the partition of  $[0, NT]$  into  $m$  equal parts of size  $\Delta t = t_{i+1} - t_i$ , where  $t_1 = 0$  and  $t_{i+1} - t_i = NTm^{-1}$  for all  $i$ . Near the end of the proof, we will choose  $m$  in terms of  $N = 1/\epsilon$ . Let  $b(t) = \max\{i : iNTm^{-1} \leq Nt\}$ . We add and subtract the

corresponding Riemann sum and use the Cauchy-Schwarz inequality to obtain

$$\begin{aligned}
& E \left[ \sup_{0 \leq t \leq T} \left| \int_0^t \left( k_j \epsilon(y_u^\epsilon)_j (y_{u-c_p}^\epsilon)_\ell - \Psi_{j\ell p} \right) du \right|^2 \right] \\
& \leq 2E \left[ \sup_{0 \leq t \leq T} \left| \frac{1}{N} \left( \int_0^{Nt} \left( k_j(z_r)_j (z_{r-c_p})_\ell - \Psi_{j\ell p} \right) dr \right. \right. \right. \\
& \quad \left. \left. \left. - \frac{NT}{m} \sum_{i=1}^{b(t)} \left( k_j(z_{t_i})_j (z_{t_i-c_p})_\ell - \Psi_{j\ell p} \right) \right) \right|^2 \right] \quad (35) \\
& \quad + \frac{2}{N^2} E \left[ \sup_{0 \leq t \leq T} \left| \frac{NT}{m} \sum_{i=1}^{b(t)} \left( k_j(z_{t_i})_j (z_{t_i-c_p})_\ell - \Psi_{j\ell p} \right) \right|^2 \right]
\end{aligned}$$

Each of the two terms on the right hand side converges to zero for a different reason. The first term goes to zero because as  $m$  increases the Riemann sum better approximates the integral. The second term goes to zero because the sum grows like  $\sqrt{m}$  since it is a sum of on the order of  $m$  weakly correlated, mean zero random variables.

We will start with the first term. First, writing

$$\frac{NT}{m} \left( k_j(z_{t_i})_j (z_{t_i-c_p})_\ell - \Psi_{j\ell p} \right) = \int_{t_i}^{t_{i+1}} \left( k_j(z_t)_j (z_{t-c_p})_\ell - \Psi_{j\ell p} \right) dt$$

we have

$$\begin{aligned}
& E \left[ \sup_{0 \leq t \leq T} \left| \frac{1}{N} \int_0^{Nt} \left( k_j(z_r)_j (z_{r-c_p})_\ell - \Psi_{j\ell p} \right) dr - \frac{T}{m} \sum_{i=1}^{b(t)} \left( k_j(z_{t_i})_j (z_{t_i-c_p})_\ell - \Psi_{j\ell p} \right) \right|^2 \right] \\
& = E \left[ \sup_{0 \leq t \leq T} \left| \frac{1}{N} \sum_{i=1}^{b(t)} \int_{t_i}^{t_{i+1}} \left[ \left( k_j(z_r)_j (z_{r-c_p})_\ell - \Psi_{j\ell p} \right) - \left( k_j(z_{t_i})_j (z_{t_i-c_p})_\ell - \Psi_{j\ell p} \right) \right] dr \right. \right. \\
& \quad \left. \left. + \frac{1}{N} \int_{t_{b(t)}}^{Nt} \left( k_j(z_r)_j (z_{r-c_p})_\ell - \Psi_{j\ell p} \right) dr \right|^2 \right] \\
& \leq 2E \left[ \sup_{0 \leq t \leq T} \left| \frac{1}{N} \sum_{i=1}^{b(t)} \int_{t_i}^{t_{i+1}} \left[ \left( k_j(z_r)_j (z_{r-c_p})_\ell - \Psi_{j\ell p} \right) - \left( k_j(z_{t_i})_j (z_{t_i-c_p})_\ell - \Psi_{j\ell p} \right) \right] dr \right|^2 \right] \\
& \quad + 2E \left[ \sup_{0 \leq t \leq T} \left| \frac{1}{N} \int_{t_{b(t)}}^{Nt} \left( k_j(z_r)_j (z_{r-c_p})_\ell - \Psi_{j\ell p} \right) dr \right|^2 \right]. \quad (36)
\end{aligned}$$

We first bound the first term on the right-hand side of (36). By the Cauchy–Schwarz inequality

$$\begin{aligned}
& E \left[ \sup_{0 \leq t \leq T} \left| \frac{1}{N} \sum_{i=1}^{b(t)} \int_{t_i}^{t_{i+1}} \left[ \left( k_j(z_r)_j(z_{r-c_p})_\ell - \Psi_{j\ell p} \right) - \left( k_j(z_{t_i})_j(z_{t_i-c_p})_\ell - \Psi_{j\ell p} \right) \right] dr \right|^2 \right] \\
& \leq E \left[ \left( \frac{1}{N} \sum_{i=1}^{b(T)} \int_{t_i}^{t_{i+1}} \left| \left( k_j(z_r)_j(z_{r-c_p})_\ell - \Psi_{j\ell p} \right) - \left( k_j(z_{t_i})_j(z_{t_i-c_p})_\ell - \Psi_{j\ell p} \right) \right| dr \right)^2 \right] \\
& \leq \frac{1}{N^2} b(T) \Delta t \sum_{i=1}^{b(T)} \int_{t_i}^{t_{i+1}} E \left[ \left( \left( k_j(z_r)_j(z_{r-c_p})_\ell - \Psi_{j\ell p} \right) - \left( k_j(z_{t_i})_j(z_{t_i-c_p})_\ell - \Psi_{j\ell p} \right) \right)^2 \right] dr.
\end{aligned}$$

Note that  $b(T) = m$ . We compute the expectation that makes up the integrand. First we expand the square:

$$\begin{aligned}
& E \left[ \left( \left( k_j(z_r)_j(z_{r-c_p})_\ell - \Psi_{j\ell p} \right) - \left( k_j(z_{t_i})_j(z_{t_i-c_p})_\ell - \Psi_{j\ell p} \right) \right)^2 \right] \\
& = E \left[ \left( k_j(z_r)_j(z_{r-c_p})_\ell - \Psi_{j\ell p} \right)^2 \right] + E \left[ \left( k_j(z_{t_i})_j(z_{t_i-c_p})_\ell - \Psi_{j\ell p} \right)^2 \right] \\
& \quad - 2E \left[ \left( k_j(z_r)_j(z_{r-c_p})_\ell - \Psi_{j\ell p} \right) \left( k_j(z_{t_i})_j(z_{t_i-c_p})_\ell - \Psi_{j\ell p} \right) \right].
\end{aligned}$$

By using Wick's Theorem [7] to compute each expectation, we obtain

$$\begin{aligned}
& E \left[ \left( \left( k_j(z_r)_j(z_{r-c_p})_\ell - \Psi_{j\ell p} \right) - \left( k_j(z_{t_i})_j(z_{t_i-c_p})_\ell - \Psi_{j\ell p} \right) \right)^2 \right] \\
& = \frac{k_j}{2k_\ell} \left( 1 - e^{-\frac{|r-t_i|}{k_j}} e^{-\frac{|r-t_i|}{k_\ell}} \right) + \frac{\delta_{j\ell}}{2} \left( e^{-2\frac{c_p}{k_j}} - e^{-\frac{|r-t_i+c_p|}{k_j}} e^{-\frac{|r-t_i-c_p|}{k_j}} \right).
\end{aligned}$$

Note that  $r \geq t_i$  and  $r - t_i \leq \Delta t$ , so that for  $\Delta t < \min_p c_p$ ,

$$-\frac{|r-t_i+c_p|}{k_j} - \frac{|r-t_i-c_p|}{k_j} = -\frac{r-t_i+c_p}{k_j} + \frac{r-t_i-c_p}{k_j} = -2\frac{c_p}{k_j}.$$

Therefore, letting  $C_1 = \max_{j,\ell} \frac{k_j}{2k_\ell}$  and  $C_2 = \max_{j,\ell} \frac{k_j+k_\ell}{k_j k_\ell}$ , we have, for  $\Delta t = \frac{NT}{m}$  sufficiently small,

$$E \left[ \left( \left( k_j(z_r)_j(z_{r-c_p})_\ell - \Psi_{j\ell p} \right) - \left( k_j(z_{t_i})_j(z_{t_i-c_p})_\ell - \Psi_{j\ell p} \right) \right)^2 \right] \leq C_1 \left( 1 - e^{-C_2 \frac{NT}{m}} \right).$$

Thus, for  $\frac{N}{m}$  sufficiently small, we have the following bound for the first term in (36):

$$\begin{aligned}
& 2E \left[ \sup_{0 \leq t \leq T} \left| \frac{1}{N} \sum_{i=1}^{b(t)} \int_{t_i}^{t_{i+1}} \left[ \left( k_j(z_r)_j(z_{r-c_p})_\ell - \Psi_{j\ell p} \right) - \left( k_j(z_{t_i})_j(z_{t_i-c_p})_\ell - \Psi_{j\ell p} \right) \right] dr \right|^2 \right] \\
& \leq \frac{2}{N^2} m \Delta t \sum_{i=1}^m \int_{t_i}^{t_{i+1}} E \left[ \left( \left( k_j(z_r)_j(z_{r-c_p})_\ell - \Psi_{j\ell p} \right) - \left( k_j(z_{t_i})_j(z_{t_i-c_p})_\ell - \Psi_{j\ell p} \right) \right)^2 \right] dr \\
& \leq \frac{2}{N^2} m^2 (\Delta t)^2 C_1 \left( 1 - e^{-C_2 \frac{NT}{m}} \right) \\
& = 2C_1 T^2 \left( 1 - e^{-C_2 \frac{NT}{m}} \right). \tag{37}
\end{aligned}$$

We now bound the second term in (36):

$$\begin{aligned}
& 2E \left[ \sup_{0 \leq t \leq T} \left| \frac{1}{N} \int_{t_{b(t)}}^{Nt} \left( k_j(z_r)_j(z_{r-c_p})_\ell - \Psi_{j\ell p} \right) dr \right|^2 \right] \\
& \leq \frac{2}{N^2} E \left[ \max_{1 \leq i \leq m} \left( \int_{t_i}^{t_{i+1}} \left| k_j(z_r)_j(z_{r-c_p})_\ell - \Psi_{j\ell p} \right| dr \right)^2 \right] \\
& \leq \frac{2}{N^2} E \left[ \sum_{i=1}^m \left( \int_{t_i}^{t_{i+1}} \left| k_j(z_r)_j(z_{r-c_p})_\ell - \Psi_{j\ell p} \right| dr \right)^2 \right] \\
& \leq \frac{2}{N^2} C_3 m (\Delta t)^2 \\
& = \frac{2C_3 T^2}{m}. \tag{38}
\end{aligned}$$

Together, (37) and (38) give a bound for the first term in (35). We now turn our attention to the second term in (35). First, we note

$$\begin{aligned}
& \frac{1}{N^2} E \left[ \sup_{0 \leq t \leq T} \left| \frac{NT}{m} \sum_{i=1}^{b(t)} \left( k_j(z_{t_i})_j(z_{t_i-c_p})_\ell - \Psi_{j\ell p} \right) \right|^2 \right] \\
& = E \left[ \max_{0 \leq n \leq m} \left| \frac{T}{m} \sum_{i=1}^n \left( k_j(z_{t_i})_j(z_{t_i-c_p})_\ell - \Psi_{j\ell p} \right) \right|^2 \right].
\end{aligned}$$

Consider the Riemann sum

$$\frac{NT}{m} \sum_{i=1}^m \left( k_j(z_{t_i})_j(z_{t_i-c_p})_\ell - \Psi_{j\ell p} \right).$$

Then

$$\begin{aligned}
& E \left[ \left( \frac{NT}{m} \sum_{i=1}^m \left( k_j(z_{t_i})_j(z_{t_i-c_p})_\ell - \Psi_{j\ell p} \right) \right)^2 \right] \\
&= E \left[ \sum_{i=1}^m \frac{NT}{m} \sum_{q=1}^m \frac{NT}{m} \left( k_j(z_{t_i})_j(z_{t_i-c_p})_\ell - \Psi_{j\ell p} \right) \left( k_j(z_{t_q})_j(z_{t_q-c_p})_\ell - \Psi_{j\ell p} \right) \right] \\
&= \sum_{i=1}^m \frac{NT}{m} \sum_{q=1}^m \frac{NT}{m} E \left[ k_j^2(z_{t_i})_j(z_{t_i-c_p})_\ell(z_{t_q})_j(z_{t_q-c_p})_\ell - \Psi_{j\ell p} k_j(z_{t_i})_j(z_{t_i-c_p})_\ell \right. \\
&\quad \left. - \Psi_{j\ell p} k_j(z_{t_q})_j(z_{t_q-c_p})_\ell + (\Psi_{j\ell p})^2 \right] \\
&= \sum_{i=1}^m \frac{NT}{m} \sum_{q=1}^m \frac{NT}{m} \left[ \frac{k_j}{4k_\ell} e^{-\frac{|t_i-t_q|}{k_j}} e^{-\frac{|t_i-t_q|}{k_\ell}} + \frac{\delta_{j\ell}}{4} e^{-\frac{|t_i-t_q+c_p|}{k_j}} e^{-\frac{|t_i-t_q-c_p|}{k_j}} \right] \\
&\leq C_4 NT
\end{aligned}$$

where  $C_4$  does not depend on  $N, m$ , or  $T$ , by comparison with the corresponding integral (recall that  $\frac{NT}{m} = t_{i+1} - t_i$ ).

Thus

$$E \left[ \left| \frac{NT}{m} \sum_{i=1}^m \left( k_j(z_{t_i})_j(z_{t_i-c_p})_\ell - \Psi_{j\ell p} \right) \right| \right] \leq C_5 \sqrt{NT}$$

which implies

$$E \left[ \left| \sum_{i=1}^m \left( k_j(z_{t_i})_j(z_{t_i-c_p})_\ell - \Psi_{j\ell p} \right) \right| \right] \leq C_5 \frac{m}{\sqrt{NT}}.$$

Applying the last bound to partial sums  $\sum_{i=1}^{2^j}$ , with  $NT$  replaced by  $\frac{2^j}{m} NT$ , we obtain

$$\Delta_r = \sum_{j=0}^{r-1} \left| \frac{E[S_{2^j}]}{2^{\frac{j}{2}}} \right| \leq C_5 \sum_{j=0}^{r-1} \frac{2^j}{\sqrt{\frac{2^j}{m} NT}} 2^{-\frac{j}{2}} = C_5 r \sqrt{\frac{m}{NT}}$$

so that by Lemma 6

$$\begin{aligned}
& E \left[ \max_{1 \leq n \leq m} \left( \sum_{i=1}^n \left( k_j(z_{t_i})_j(z_{t_i-c_p})_\ell - \Psi_{j\ell p} \right) \right)^2 \right] \\
&\leq m \left( 2C_6 + (1 + \sqrt{2}) C_5 \sqrt{\frac{m}{NT}} \log_2(m) \right)^2
\end{aligned}$$



where  $C_6 = \sqrt{E \left[ \left( k_j(z_{t_i})_j(z_{t_i-c_p})_\ell - \Psi_{j\ell p} \right)^2 \right]}$ . Thus,

$$\begin{aligned} E \left[ \max_{1 \leq n \leq m} \left( \frac{T}{m} \sum_{i=1}^n \left( k_j(z_{t_i})_j(z_{t_i-c_p})_\ell - \Psi_{j\ell p} \right) \right)^2 \right] \\ \leq \frac{T^2}{m} \left( 2C_6 + (1 + \sqrt{2}) C_5 \sqrt{\frac{m}{NT}} \log_2(m) \right)^2. \end{aligned} \quad (39)$$

Putting together (37), (38), and (39), we have, for  $\frac{N}{m}$  sufficiently small,

$$\begin{aligned} E \left[ \sup_{0 \leq t \leq T} \left| \int_0^t \left( k_j \epsilon(y_u)_j(y_{u-c_p\epsilon})_\ell - \Psi_{j\ell p} \right) du \right|^2 \right] \\ \leq 2C_1 T^2 \left( 1 - e^{-C_2 \frac{NT}{m}} \right) + \frac{2T^2}{m} \left( 2C_6 + (1 + \sqrt{2}) C_5 \sqrt{\frac{m}{NT}} \log_2(m) \right)^2 + \frac{2C_3 T^2}{m}. \end{aligned}$$

Setting  $m = N^2$  proves (34) and thus (32) and (33).

### 3.3.3 Condition 1

Now we show that  $\mathbf{H}^\epsilon$  satisfies Condition 1. To do so, we must show that for every  $j, \ell$ , and  $p$ ,

$$\left| \int_0^t k_j \epsilon(y_s)_j(y_{s-c_p\epsilon})_\ell - \frac{1}{2} e^{-\frac{c_p}{k_j}} ds \right|$$

and

$$\left| \int_0^t k_j \epsilon(y_s)_j(y_{s-c_p\epsilon})_\ell ds \right|$$

are stochastically bounded. This follows from Lemma 3 by Chebyshev inequality, since the second moments of the integrands are bounded by a constant, independent of  $\epsilon$ .

We now complete the proof of Theorem 1. From Section 3.3.1,  $\mathbf{U}^\epsilon \rightarrow \mathbf{0}$  in probability with respect to  $C([0, T], \mathbb{R}^m)$ . From Lemma 5,  $\mathbf{H}^\epsilon \rightarrow \mathbf{H}$  in  $L^2$  with respect to  $C([0, T], \mathbb{R}^{(1+n+1+mn^2)})$ . Therefore,  $(\mathbf{U}^\epsilon, \mathbf{H}^\epsilon) \rightarrow (\mathbf{0}, \mathbf{H})$  in probability with respect to  $C([0, T], \mathbb{R}^m \times \mathbb{R}^{(1+n+1+mn^2)})$ . Furthermore  $\mathbf{H}^\epsilon$  satisfies Condition 1 by Lemma 3. Thus, by Lemma 1,  $\mathbf{x}^\epsilon \rightarrow \mathbf{x}$  in probability with respect to  $C([0, T], \mathbb{R}^m)$ .  $\square$

## 4 Discussion

We have derived the limiting equation for a general stochastic differential delay equation with multiplicative colored noise as the time delays and correlation

times of the noises go to zero at the same rate. As a result of the dependence of the noise coefficients on the state of the system (multiplicative noise), a *noise-induced drift* appears in the limiting equation. This result is useful for applications as the limiting SDE could provide a model that is easier to analyze than the original equation and at the same time still accounts for the effects of the time delays through the coefficients of the noise-induced drift.

The noise-induced drift has a form analogous to that of the *Stratonovich correction* to the Itô equation with noise term  $\sigma(\mathbf{x}_t)d\mathbf{W}_t$ . That is, the noise-induced drift is a linear combination of the terms  $\sigma_{pj}(\mathbf{x}_t)\frac{\partial\sigma_{ij}(\mathbf{x}_t)}{\partial x_p}$ . The coefficients of this linear combination in the limiting equation (3) are

$$\alpha_{jp} = \frac{1}{2}e^{-\frac{\delta_p}{\tau_j}}, \quad (40)$$

whereas the coefficients of the Stratonovich correction would all be equal to  $\frac{1}{2}$ . Thus, as explained in [10, 16], one can interpret the terms of the noise-induced drift as representing different stochastic integration conventions.

The limiting equation that we have derived here is a more accurate approximation of the delayed system than the limiting equation derived in [10, 16]. In particular, in [10, 16], equation (1) was Taylor expanded to first-order in the time delay and the limiting equation corresponding to the expanded equation was derived. It was shown that this limiting equation contains a noise-induced drift  $\tilde{\mathbf{S}}(\mathbf{x}_t)$  defined componentwise as

$$\tilde{S}_i(\mathbf{x}) = \sum_{p,j} \frac{1}{2} \left(1 + \frac{\delta_p}{\tau_j}\right)^{-1} \sigma_{pj}(\mathbf{x}) \frac{\partial\sigma_{ij}(\mathbf{x})}{\partial x_p}. \quad (41)$$

The coefficient  $\frac{1}{2} \left(1 + \frac{\delta_p}{\tau_j}\right)^{-1}$  is a first-order Taylor expansion in the parameter  $\frac{\delta_p}{\tau_j}$  of the coefficient  $\frac{1}{2}e^{-\frac{\delta_p}{\tau_j}}$  in (4) in the sense that  $\left(1 + \frac{\delta_p}{\tau_j}\right)^{-1}$  is obtained when one expands the denominator of  $1/e^{\delta_p/\tau_j}$  to first-order in  $\frac{\delta_p}{\tau_j}$ . Thus, while the two limiting equations are close when all the ratios  $\frac{\delta_p}{\tau_j}$  are small, the limiting equation derived here is overall a better approximation of the delayed system.

## 5 Conclusion

We have proven a result concerning convergence of the solution of a general SDDE driven by state-dependent colored noise to the solution of an SDE driven by white noise. The main theorem (Theorem 1) was proven using direct analysis of the SDDE without Taylor expansion and thus it improves upon a result from previous works. The resulting limiting equation (3) was compared to the previous results in [10, 16]. The noise-induced drift derived there was seen to be a first-order expansion in the ratios  $\delta_p/\tau_j$  of the one found here. The limiting equation derived here can be used as an approximation to study the dynamics of real systems modeled by the SDDE.

## Acknowledgements

A.M. and J.W. were partially supported by the NSF grants DMS 1009508 and DMS 0623941.

## References

- [1] G. Da Prato, S. Kwapień, and J. Zabczyk. Regularity of solutions of linear stochastic equations in Hilbert spaces. *Stochastics: An International Journal of Probability and Stochastic Processes*, 23(1):1–23, 1988.
- [2] G. Da Prato and J. Zabczyk. *Stochastic equations in infinite dimensions*, volume 44 of *Encyclopedia Math. Appl.* Cambridge University Press, Cambridge, UK, 1992.
- [3] M. Freidlin. Some remarks on the Smoluchowski-Kramers approximation. *J. Stat. Phys.*, 117:617–634, 2004.
- [4] C. W. Gardiner. *Stochastic methods*. Springer, Berlin, 2009.
- [5] S. Hottovy, A. McDaniel, G. Volpe, and J. Wehr. The Smoluchowski-Kramers limit of stochastic differential equations with arbitrary state-dependent friction. *Communications in Mathematical Physics*, 336(3):1259–1283, 2014.
- [6] S. Hottovy, G. Volpe, and J. Wehr. Thermophoresis of Brownian particles driven by coloured noise. *EPL (Europhys. Lett.)*, 99:60002, 2012.
- [7] S. Janson. *Gaussian Hilbert spaces*, volume 129 of *Cambridge Tracts in Mathematics*. Cambridge University Press, Cambridge, 1997.
- [8] R. Kupferman, G. A. Pavliotis, and A. M. Stuart. Itô versus Stratonovich white-noise limits for systems with inertia and colored multiplicative noise. *Phys. Rev. E*, 70:036120, 2004.
- [9] T. G. Kurtz and P. Protter. Weak limit theorems for stochastic integrals and stochastic differential equations. *Ann. Probab.*, 19:1035–1070, 1991.
- [10] A. McDaniel, O. Duman, G. Volpe, and J. Wehr. An SDE approximation for stochastic differential delay equations with colored state-dependent noise. *arXiv preprint arXiv:1406.7287*, 2014.
- [11] E. Nelson. *Dynamical theories of Brownian motion*. Princeton University Press, Princeton, N.J., 1967.
- [12] B. Øksendal. *Stochastic differential equations*. Universitext. Springer-Verlag, Berlin, sixth edition, 2003. An introduction with applications.

- [13] G. A. Pavliotis and A. M. Stuart. Analysis of white noise limits for stochastic systems with two fast relaxation times. *Multiscale Model. Simul.*, 4(1):1–35 (electronic), 2005.
- [14] G. A. Pavliotis and A. M. Stuart. *Multiscale Methods: Averaging and Homogenization*, volume 53 of *Texts in Applied Mathematics*. Springer, New York, 2008.
- [15] M. Peligrad and S. Utev. A new maximal inequality and invariance principle for stationary sequences. *Ann. Probab.*, 33(2):798–815, 2005.
- [16] G. Pesce, A. McDaniel, S. Hottovy, J. Wehr, and G. Volpe. Stratonovich-to-Itô transition in noisy systems with multiplicative feedback. *Nature Communications*, 4, 2013.
- [17] P. E. Protter. *Stochastic integration and differential equations*, volume 21 of *Stochastic Modelling and Applied Probability*. Springer-Verlag, Berlin, 2005. Second edition. Version 2.1, Corrected third printing.
- [18] D. Revuz and M. Yor. *Continuous martingales and Brownian motion*, volume 293 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, third edition, 1999.